

POLYTOPES FROM SUBGRAPH STATISTICS

ALEXANDER ENGSTRÖM AND PATRIK NORÉN

ABSTRACT. Polytopes from subgraph statistics are important in applications and conjectures and theorems in extremal graph theory can be stated as properties of them. We have studied them with a view towards applications by inscribing large explicit polytopes and semi-algebraic sets when the facet descriptions are intractable. The semi-algebraic sets called curvy zonotopes are introduced and studied using graph limits. From both volume calculations and algebraic descriptions we find several interesting conjectures.

1. INTRODUCTION

In this paper we study polytopes from subgraph statistics. The vertices of these polytopes are given by the relative proportions of subgraphs of different types. We got interested in studying the polytopes from subgraph statistics after several questions were raised about them by Rinaldo, Fienberg and Zhou [21]. They investigated maximum likelihood estimation for exponential random graph models and realized that its behavior is closely linked to the geometry of these polytopes. The subgraphs counted are usually determined by the applications of the model, and in the social sciences small graphs as stars and triangles are common [22], but we make no restrictions of that type.

Our results and methods are from graph theory and discrete geometry, but we have made an effort to address directions that are important for applications. For example to understand when the polytopes have the expected dimensions and how to approximate them with similar polytopes or semi-algebraic sets with easy explicit descriptions when their facet structures are not attainable.

On the track to these descriptions we have found several, for us, unexpected results and conjectures about both enumerative and geometric combinatorics.

For the experts who wants to skip ahead, we want to clarify a notational difference between mathematical communities: Our focus is on finite graphs and finite dimensional polytopes close to applications, even if we use limits and infinite objects as technical tools in some of the later proofs. Therefore the t -function counts ordinary honest subgraphs and not graph homomorphisms, and these notions are crucially different in the finite setting before the limit.

1.1. A short overview of the paper. In Section 2 we define the polytope from subgraphs statistics and its lattice version. We explain basic properties of their facet structures and how different polytopes and Ehrhart polynomials are related to each other. A complete facet description would solve many open problems in extremal graph theory, so to get any understanding of these polytopes we inscribe more well-studied polytopes and semi-algebraic sets in them.

The spine of our polytopes is defined in Section 3. The convex hull of a finite number of points on the spine is a cyclic polytope inscribed in the polytope of subgraph statistics, and the convex hull of the spine is a semi-algebraic set. In Section 3.1 we calculate the volumes of the inscribed sets and find both explicit formulas and integrals connected to the Selberg integral formula. The dual of the convex hull of the spine and the polytope provides a method to find certificates that polynomials are non-negative.

The convex hull of the spine never fills up all of the polytope, and can even be of lower dimensions. In Section 4 we introduce the curvy zonotopes. They are semi-algebraic sets with explicit descriptions that are of the top dimension. Alternatively the spine could have been defined as the expected values of graphs from the Erdős-Renyi graph model with different model parameters. We get the curvy zonotopes as expected values of particular exchangeable graph models.

In Section 5 we show that the curvy zonotopes are not only of the right dimension, but that they can be chosen to get the volume arbitrary close to the polytope that they are inscribed in. The proofs relies on the theory of graph limits and Szemerédi regularity. Finally in Section 6 we study the limit case of counting complete subgraphs and conjecture that the limit objects essentially are cyclic polytopes with infinite many vertices.

2. BASIC PROPERTIES OF POLYTOPES OF SUBGRAPH STATISTICS

In this section we give proper definitions of the objects of our study, and provide some first results to describe them. We follow standard notation in graph theory, as in for example Diestel [7].

The number of subgraphs of G isomorphic to F is $t^L(F, G)$. Later the letter L indicates that we work with lattice polytopes. The F -subgraph density in G is defined as

$$t(F, G) = \frac{t^L(F, G)}{t^L(F, K_{|G|})},$$

except when F has more vertices than G , and then it is 0.

Definition 2.1. Let $\mathbf{F} = (F_1, \dots, F_d)$ be a vector of graphs and G a graph. The vector valued subgraph statistics are

$$t(\mathbf{F}, G) = (t(F_1, G), \dots, t(F_d, G))$$

and

$$t^L(\mathbf{F}, G) = (t^L(F_1, G), \dots, t^L(F_d, G)).$$

Definition 2.2. Let \mathbf{F} be a vector of graphs and n a positive integer. The *polytope from subgraph statistics* $P_{\mathbf{F};n}$ and its lattice version $P_{\mathbf{F};n}^L$ are defined as

$$P_{\mathbf{F};n} = \text{conv} \{t(\mathbf{F}, G) \mid G \text{ is a graph on } n \text{ vertices}\}$$

and

$$P_{\mathbf{F};n}^L = \text{conv} \{t^L(\mathbf{F}, G) \mid G \text{ is a graph on } n \text{ vertices}\}.$$

Example 2.3. The polytope $P_{(K_3, C_4, K_4 \setminus e);6}$ is drawn in Figure 1, and in Figure 2 is a combinatorial representation of its vertices and edges.

If larger examples of polytopes from subgraph statistics looks anything like in Figures 1 and 2, then it will be very difficult to give an explicit facet description. And indeed many hard theorems and conjectures in extremal graph theory can be rephrased as questions about these polytopes, making a complete facet description probably impossible in general. In Figure 2 we tabulated the vertices by the actual subgraph counts and not the proportions $t(F, G)$. This defines the lattice polytope $P_{\mathbf{F};n}^L$, a rescaling of $P_{\mathbf{F};n}$. Several

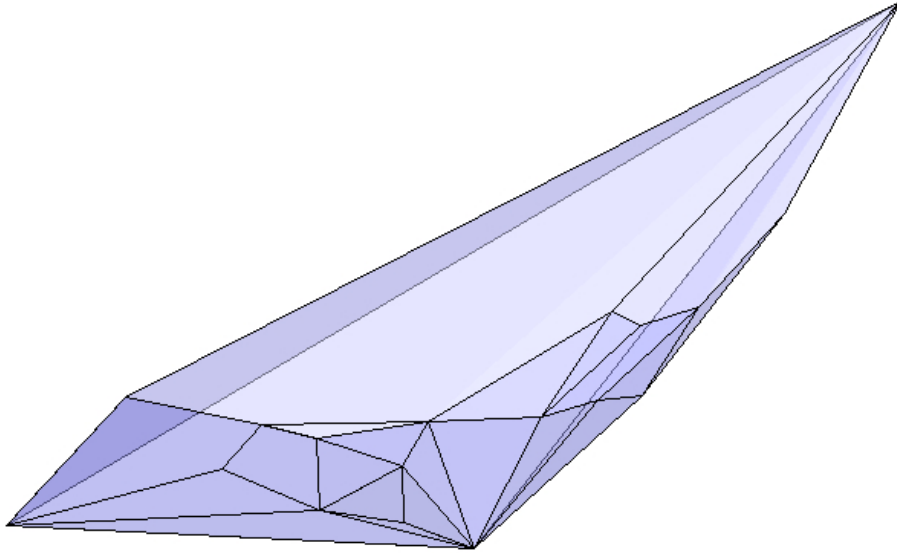


FIGURE 1. The polytope $P_{(K_3, C_4, K_4 \setminus e);6}$.

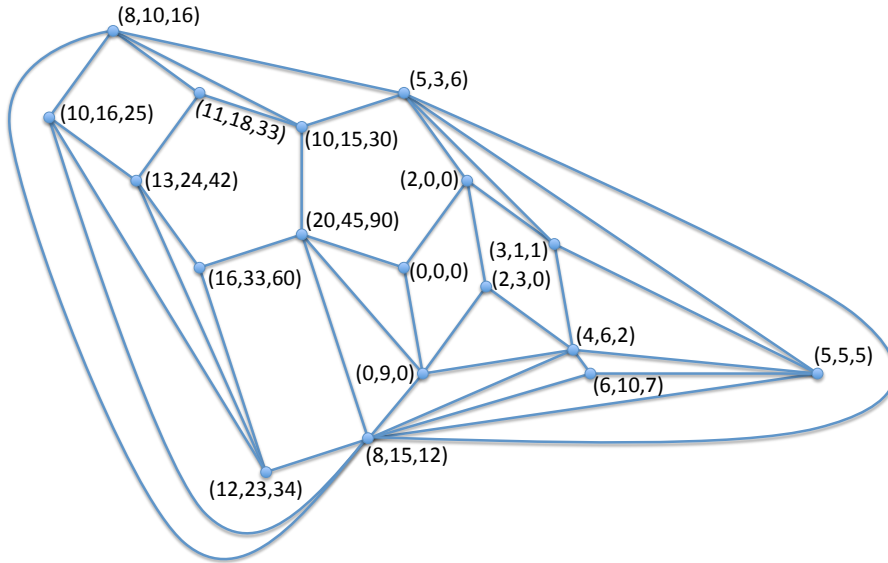


FIGURE 2. A combinatorial representation of the vertices and edges of $P_{(K_3, C_4, K_4 \setminus e);6}$, indexed by the actual subgraph counts.

graphs could have the same subgraph statistics, and even if $t(\mathbf{F}, G_1)$ and $t(\mathbf{F}, G_2)$ are different vertices on the same facet, it is not necessary that G_1 and G_2 are related in any sense, for example as subgraphs. This is illustrated in Figure 3.

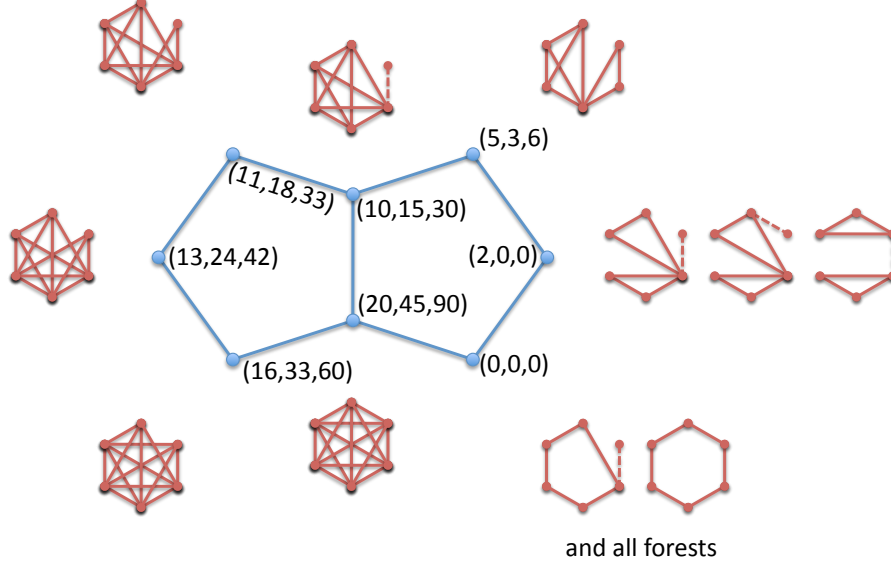


FIGURE 3. The graphs underlying the statistics of a piece of the polytope in Figures 1 and 2. Dotted graph edges could be included or not. Recall that the subgraphs counted are $(K_3, C_4, K_4 \setminus e)$.

Before embarking on more general results about the polytopes from subgraph statistics, we point out some easy facts.

Proposition 2.4. *Let $\mathbf{F} = (F_1, \dots, F_d)$ be a vector of graphs with edges of order at least n , none of them a subgraph of another. Then the inequalities $x_i \geq 0$ are facet defining for $P_{\mathbf{F};n}$.*

Proof. The number of subgraphs is non-negative and the inequality $x_i \geq 0$ always defines a face, the question is if it is a facet. Let G_j be the graph given by the disjoint union of F_j and enough isolated vertices to get n vertices in total. The vector $t(\mathbf{F}, G_j)$ is non-zero exactly in the component j . On the hyperplane defined by $x_i = 0$ is a dimension $d - 1$ simplex defined by the points $\mathbf{0}$ and $t(\mathbf{F}, G_j)$ for $i \neq j$. \square

Lemma 2.5. *If F and G are graphs then $t^L(F, G) = \frac{|F|!}{\text{Aut}(F)} \binom{|G|}{|F|} t(F, G)$.*

Proof. If $t(F, G) = t^L(F, G) = 0$ it is true. Otherwise use that $t^L(F, K_{|G|}) = \frac{|F|!}{\text{Aut}(F)} \binom{|G|}{|F|}$. \square

Lemma 2.6. *If F and G are graphs and n an integer then*

$$\binom{|G| - |F|}{n - |F|} t^L(F, G) = \sum_{U \in \binom{V(G)}{n}} t^L(F, G[U]).$$

Proof. If not $|F| \leq n \leq |G|$ then both sides are zero. Otherwise, consider a particular copy of F in G . On the right hand side it is counted for every subset U of size n containing the vertex set of that copy of F . The set U contains $n - |F|$ elements to be chosen among the $|G| - |F|$ vertices outside that particular copy of F . \square

Proposition 2.7. *Let \mathbf{F} be a vector of graphs of order at most n . If $n \leq n' \leq n''$ then $P_{\mathbf{F};n'} \supseteq P_{\mathbf{F};n''}$.*

Proof. Let G be a graph on n'' vertices and let F be any graph in \mathbf{F} . First combine Lemma 2.5 and Lemma 2.6 to get

$$t(F, G) = \frac{\binom{n'}{|F|}}{\binom{|G|}{|F|} \binom{|G|-|F|}{n'-|F|}} \sum_{U \in \binom{V(G)}{n'}} t(F, G[U]).$$

By expanding and simplifying, one gets that $\frac{\binom{n'}{|F|}}{\binom{|G|}{|F|} \binom{|G|-|F|}{n'-|F|}} = \binom{|G|}{n'}^{-1}$. Collected into vectors this is

$$t(\mathbf{F}, G) = \binom{|G|}{n'}^{-1} \sum_{U \in \binom{V(G)}{n'}} t(\mathbf{F}, G[U])$$

and all vertices of $P_{\mathbf{F};n''}$ are in simplices spanned by points in $P_{\mathbf{F};n'}$. \square

For any number k , a polytope P can be inflated to $kP = \{kp | p \in P\}$. For a lattice polytope P and positive integers k , the number of lattice points in kP is the *Ehrhart polynomial* $E_P(k)$. We refer to chapter 12 of Miller and Sturmfels [17] for a proof of this, and a description of the connections to algebraic geometry. This is a translation of Proposition 2.7 into the lattice polytope setting.

Proposition 2.8. *Let $\mathbf{F} = (F_1, \dots, F_d)$ be a vector of graphs of order n . Then*

$$E_{P_{\mathbf{F};n''}}^L \left(\binom{n'}{n} k \right) \leq E_{P_{\mathbf{F};n'}}^L \left(\binom{n''}{n} k \right)$$

if $n'' \geq n' \geq n$.

Proof. Define a linear rescaling map $L_{\mathbf{F}} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ by $L_{\mathbf{F}}(\mathbf{e}_i) = \text{Aut}(F_i)^{-1} \mathbf{e}_i$. By Lemma 2.5,

$$L_{\mathbf{F}} \left(n! \binom{n''}{n} P_{\mathbf{F};n''} \right) = P_{\mathbf{F};n''}^L \text{ and } L_{\mathbf{F}} \left(n! \binom{n'}{n} P_{\mathbf{F};n'} \right) = P_{\mathbf{F};n'}^L.$$

According to Proposition 2.7 there is an inclusion of polytopes $P_{\mathbf{F};n''} \subseteq P_{\mathbf{F};n'}$ since $n'' \geq n'$. We rescale the inclusion to

$$\binom{n'}{n} \left(n! \binom{n''}{n} P_{\mathbf{F};n''} \right) \subseteq \binom{n''}{n} \left(n! \binom{n'}{n} P_{\mathbf{F};n'} \right),$$

apply the linear rescaling map and move out some scalar factors,

$$\binom{n'}{n} L_{\mathbf{F}} \left(n! \binom{n''}{n} P_{\mathbf{F};n''} \right) \subseteq \binom{n''}{n} L_{\mathbf{F}} \left(n! \binom{n'}{n} P_{\mathbf{F};n'} \right),$$

or equivalently, $\binom{n'}{n} P_{\mathbf{F};n''}^L \subseteq \binom{n''}{n} P_{\mathbf{F};n'}^L$. Counting lattice points gives the desired result. \square

In the proposition it is required that all graphs in \mathbf{F} are of the same order, and this can partially be generalized by adding isolated vertices to get graphs of the same order.

Example 2.9. If \mathbf{F} is the graph vector of the path on three vertices and the triangle, then

$$E_{P_{\mathbf{F};3}}^L \left(\binom{4}{3} k \right) = E_{P_{\mathbf{F};4}}^L \left(\binom{3}{3} k \right) = 8k^2 + 6k + 1$$

and

$$E_{P_{\mathbf{F};3}}^L \left(\binom{5}{3} k \right) = 50k^2 + 15k + 1 \geq 48k^2 + 13k + 1 = E_{P_{\mathbf{F};5}}^L \left(\binom{3}{3} k \right).$$

3. THE SPINE OF POLYTOPES

Since it's hard to understand the polytopes exactly, we inscribe more accessible polytopes and semi-algebraic sets within them. For a vector \mathbf{F} of m graphs, the *spine* is the generalized moment curve

$$\{(p^{e_1}, p^{e_2}, \dots, p^{e_m}) \mid 0 \leq p \leq 1\},$$

where e_i is the number of edges in F_i . A graph G from the Erdős-Rényi random graph model $\mathcal{G}(n, p)$ have n vertices, and edges are included independently with probability p .

Proposition 3.1. *Let $\mathbf{F} = (F_1, \dots, F_d)$ be a vector of graphs of order at most n , then the spine $\{(p^{e_1}, p^{e_2}, \dots, p^{e_d}) \mid 0 \leq p \leq 1\}$ is in $P_{\mathbf{F};n}$.*

Proof. The expected value of $t(F, G)$ with G from $\mathcal{G}(n, p)$ is p^e if F have e edges, since the edges of G are included independently with probability p . For every instance of graphs G from $\mathcal{G}(n, p)$ the vector $t(\mathbf{F}, G)$ is a point in $P_{\mathbf{F};n}$. The expected value of $t(\mathbf{F}, G)$,

$$(p^{e_1}, p^{e_2}, \dots, p^{e_m}),$$

is then also a point in $P_{\mathbf{F};n}$. □

In Figure 4 is the polytope $P_{(K_3, C_4, K_4 \setminus e);6}$ from Figure 1 drawn with its spine. The point of Proposition 3.1 is that the spine is a generalized moment curve inside $P_{\mathbf{F};n}$. The convex hull of a finite number of points on the spine is a cyclic polytope. This can be seen directly by using generalized van der Monde matrices instead of the ordinary one in Ziegler's textbook derivation of the combinatorial structure of cyclic polytopes [28]. This shows that there are cyclic polytope inscribed in $P_{\mathbf{F};n}$.

The convex hull of all of the spine is not a polytope, but its boundary can be algebraically described. In Figure 5 is the spine from Figure 4 drawn with its convex hull. Since the boundary structure of the convex hull in Figure 5 is not very clear from this angle, we include in Figure 6 the same spine with its convex hull, but from another perspective.

3.1. Volumes. Inside our polytopes we have convex hulls of generalized moment curves and their volumes bound the volumes of polytopes from subgraph statistics. For the ordinary moment curve $\{(p, p^2, \dots, p^d) \mid 0 \leq p \leq 1\}$ the volume of its convex hull was calculated by Karlin and Shapley [14]. To calculate the volume of the convex hull of the spine, and for later applications, we need the following standard approximation.

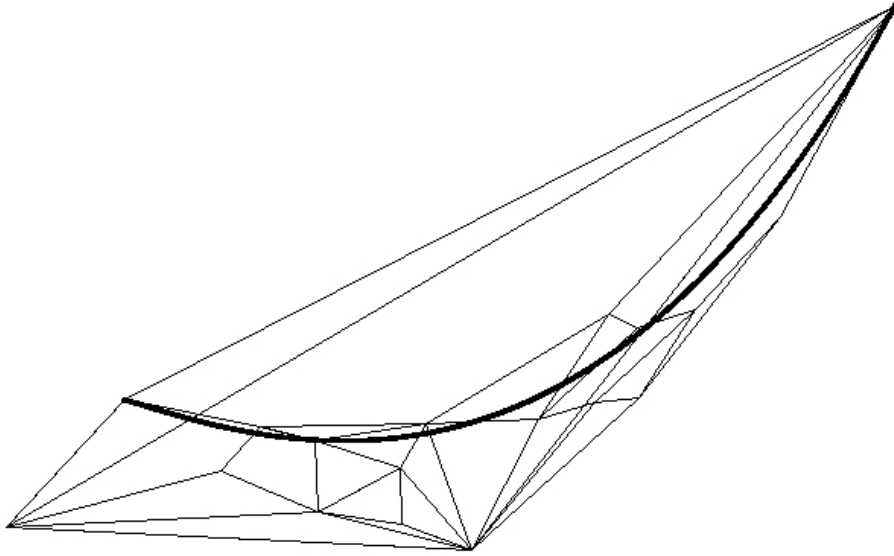


FIGURE 4. The polytope in Figure 1 with its spine.

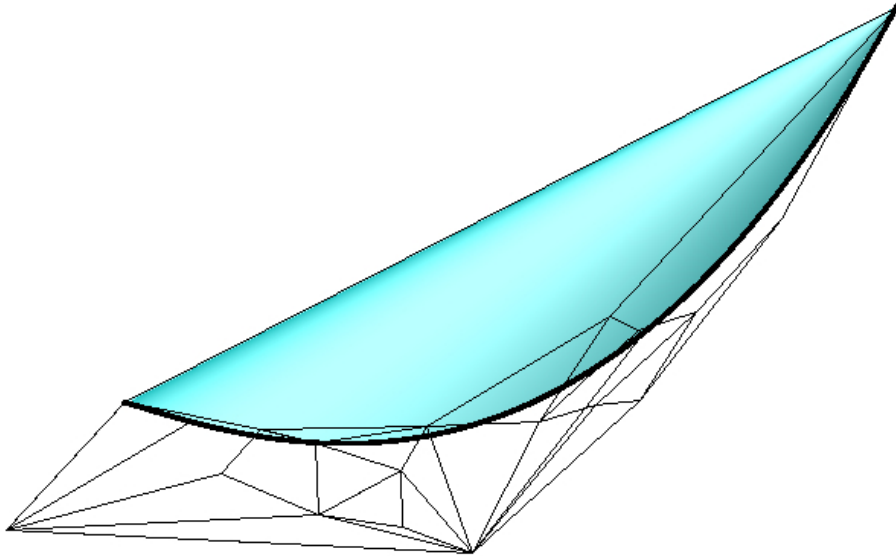


FIGURE 5. The spine in Figure 4 drawn with its convex hull.

Lemma 3.2. *Let $A \subseteq B \subseteq [0, 1]^d$ be sets so that for any $b \in B$ there is an $a \in A$ with $|a - b| \leq \varepsilon$, then*

$$\text{Vol}(\text{conv} B) - \text{Vol}(\text{conv} A) \leq k_d \varepsilon$$

where k_d only depends on the dimension d .

Proof. For $X = A, B$ and $\xi \in \mathbb{R}^d$ set

$$H_X(\xi) = \sup_{x \in X} \langle x, \xi \rangle = \sup_{x \in \text{conv} X} \langle x, \xi \rangle.$$

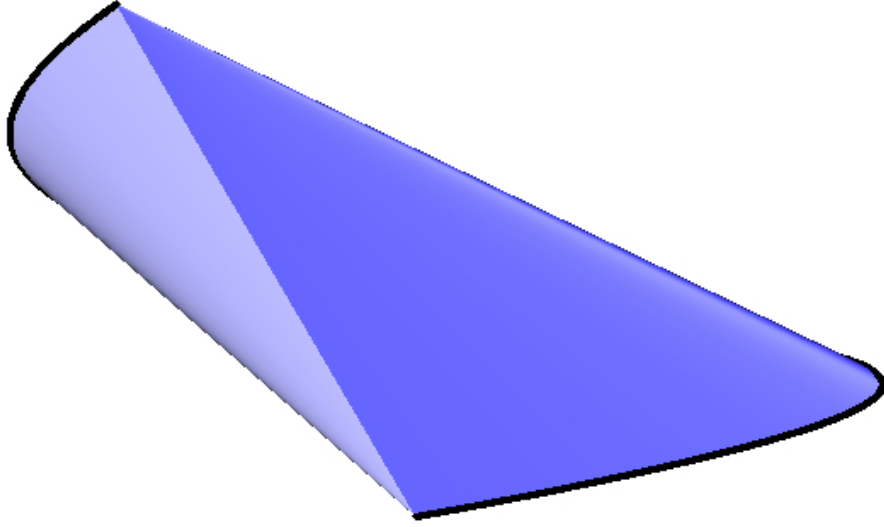


FIGURE 6. The spine with its convex hull from Figure 5 drawn from another perspective.

By assumption, there is always a point within distance ε in A from each point in B , so $H_B(\xi) - \varepsilon \leq H_A(\xi) \leq H_B(\xi)$ for all ξ . Using the setup of Hörmander [11], section 4.3, the function $H_C(\xi) = H_B(\xi) - \varepsilon$ defines a convex body $C = \{x \mid \langle x, \xi \rangle \leq H_C(\xi), \xi \in \mathbb{R}^d\} \subseteq \text{conv } A \subseteq \text{conv } B$ with

$$\text{Vol}(\text{conv } B) - \text{Vol}(C) \leq \varepsilon S_B$$

where S_B is the surface area of the closed convex hull of B . The surface area can be estimated from above with a k_d since the volume of B is at most one, for example as in Section I.8.3 of [3]. \square

Proposition 3.3. *Let $e_1 < e_2 < \dots < e_d$, and n be positive integers. Consider the cyclic polytope P on the vertices*

$$\{(x^{e_1}, x^{e_2}, \dots, x^{e_d}) \mid x = i/n, i = 0, 1, 2, \dots, n\}$$

and the convex hull C of the moment curve

$$\{(x^{e_1}, x^{e_2}, \dots, x^{e_d}) \mid 0 \leq x \leq 1\}.$$

Then

$$0 \leq \text{Vol}(C) - \text{Vol}(P) \leq \frac{k_d}{n}(e_1 + e_2 + \dots + e_d),$$

where k_d is a constant depending only on the dimension d .

Proof. Let $f : [0, 1] \rightarrow \mathbb{R}^d$ be $x \mapsto (x^{e_1}, x^{e_2}, \dots, x^{e_d})$. From the vector valued mean value theorem

$$|f(y) - f(x)| \leq |y - x| \sup_{0 \leq t \leq 1} |f'(y + t(x - y))|$$

applied with $y = \lfloor xn \rfloor / n$ and $|f'(y + t(y - x))| \leq |f'(1)| = e_1 + e_2 + \dots + e_d$, we get that any point on the moment curve is within distance

$$\varepsilon = \frac{e_1 + e_2 + \dots + e_d}{n}$$

of a vertex of the polytope P . Now use Lemma 3.2. \square

We treat Schur polynomials as they are defined by fractions of generalized Vandermonde matrices instead of using representation theory. For these elementary facts we refer to Sagan [23]. We will also use the Pfaffian, and refer to [9, 16] for definitions and basic identities.

Theorem 3.4. *Let $m > 1$ and let $\mathbf{F} = (F_1, \dots, F_{2m})$ be a vector of $2m$ distinct graphs with edges. The volume of the convex hull of the $2m$ -dimensional spine of \mathbf{F}*

$$\text{Vol}(\text{conv} \{(p^{e_1}, p^{e_2}, \dots, p^{e_{2m}}) \mid 0 \leq p \leq 1\})$$

is

$$\frac{1}{(2m)!m!} \int_{[0,1]^m} S_\lambda(x_1, x_1, x_2, x_2, \dots, x_m, x_m) \prod_{0 \leq i < j \leq m} (x_i - x_j)^4 dx_1 \cdots dx_m,$$

where $\lambda_i + 2m - i = e_i$ in the Schur polynomial S_λ , and it is assumed that all e_i are different and greater than 0.

And in the odd dimensional case $\mathbf{F} = (F_1, \dots, F_{2m+1})$ we get

$$\frac{1}{(2m+1)!m!} \int_{[0,1]^m} S_\lambda(x_1, x_1, x_2, x_2, \dots, x_m, x_m, 1) \prod_{0 \leq i < j \leq m} (x_i - x_j)^4 \prod_{0 \leq i \leq m} (1 - x_i)^2 dx_1 \cdots dx_m,$$

where $\lambda_i + 2m + 1 - i = e_i$ in the Schur polynomial S_λ , and it is assumed that all e_i are different and greater than 0.

Proof. The idea of the proof is to approximate the volume of the convex hull of the spines with the volume of cyclic polytopes. The cyclic polytopes used to approximate the volume have their vertices placed uniformly distributed on the spine. When the number of vertices of these polytopes go to infinity the volume of the polytopes converge to the volume of the convex hull of the spine.

To compute the volume of a cyclic polytope P we begin by triangulating the polytope P in a special way. The volume of the polytope P is then computed by summing the volumes of the simplexes in the triangulation. The triangulation used is constructed as follows: Pick a vertex v in P , the triangulation is the one consisting of all the simplexes spanned by a facet of P together with v . This is a triangulation of P since cyclic polytopes are simplicial. Note that the facets containing v do not contribute to the volume and so they can be ignored.

Consider the cyclic polytope P_n spanned by $\{((i/n)^{e_1}, (i/n)^{e_2}, \dots, (i/n)^{e_{2m}}) \mid 0 \leq i \leq n\}$. The polytope P_n has vertices uniformly distributed on the spine of \mathbf{F} . When n goes to infinity the volume of P_n converge to the volume of the convex hull of the spine. Choose the vertex v to be $\mathbf{0}$. By Gale's evenness condition [28] the volume of P_n is

$$\frac{1}{(2m)!} \sum_{i_1, \dots, i_m} \det \left[\left(\frac{i_k}{n} \right)^{e_j} \left(\frac{i_k + 1}{n} \right)^{e_j} \right]_{\substack{1 \leq j \leq 2m \\ 1 \leq k \leq m}}$$

where the sequences summed over are those satisfying $i_j + 1 < i_{j+1}$, $0 < i_1$ and $i_m < n$.

The determinants can be expressed in terms of Schur polynomials and so the volume of the polytope P_n is

$$\frac{1}{(2m)!} \sum_{i_1, \dots, i_{2m}} S_\lambda \left(\frac{i_1}{n}, \frac{i_1+1}{n}, \dots, \frac{i_m}{n}, \frac{i_m+1}{n} \right) \frac{1}{n^m} \prod_{1 \leq j < k \leq m} \frac{(i_k - i_j)^2 ((i_k - i_j)^2 - 1)}{n^4},$$

where the sum is over the same sequences as in the formula with determinants. This is a Riemann sum and so when n go to infinity this converges to the integral

$$\frac{1}{(2m)!} \int_I S_\lambda(x_1, x_1, \dots, x_m, x_m) \prod_{1 \leq i < j \leq m} (x_i - x_j)^4 dx_1 \cdots dx_m,$$

where $I = \{0 < x_1 < x_2 < \dots < x_m < 1\}$. Both the Schur polynomial and the product of four-powers of differences are symmetric polynomials, and we can replace I by $[0, 1]^m$ and divide by a factor $m!$ to get the integral stated in the theorem.

The integral is the volume of the convex hull of the spine according to Proposition 3.3.

The odd dimensional case is similar, but every full dimensional simplex contains the vertex $\mathbf{1}$ since the facets of a cyclic polytope in odd dimensions all contain $\mathbf{0}$ or $\mathbf{1}$. \square

In the even dimensional case there is a nice formula for the integral in Theorem 3.4.

Theorem 3.5. *Let $\lambda = (\lambda_1, \dots, \lambda_{2m})$ in the Schur polynomial S_λ , then*

$$\int_{[0,1]^m} S_\lambda(x_1, x_1, x_2, x_2, \dots, x_m, x_m) \prod_{0 \leq i < j \leq m} (x_i - x_j)^4 d\mathbf{x} = m! \prod_{1 \leq i < j \leq 2m} \frac{y_i - y_j}{y_i + y_j}$$

where $y_i = 2m + \lambda_i - i$.

In the odd dimensional case, let $\lambda = (\lambda_1, \dots, \lambda_{2m+1})$ in the Schur polynomial S_λ . Then

$$\int_{[0,1]^m} S_\lambda(x_1, x_1, x_2, x_2, \dots, x_m, x_m, 1) \prod_{0 \leq i < j \leq m} (x_i - x_j)^4 \prod_{0 \leq i \leq m} (1 - x_i)^2 d\mathbf{x} = m! \prod_{1 \leq i < j \leq 2m+1} \frac{y_i - y_j}{y_i + y_j},$$

where $y_i = 2m + 1 + \lambda_i - i$.

Proof. First we prove it in the even dimensional case. The first step of the proof is to go back to expressing the integrand as a limit of a determinant. This limit is also a nice determinant that is possible to integrate. Observe that $S_\lambda(x_1, x_1, x_2, x_2, \dots, x_m, x_m) \prod_{1 \leq i < j \leq m} (x_j - x_i)^4$ is the limit of

$$\frac{\det[x_i^{2m-j+\lambda_j}]_{1 \leq i, j \leq 2m} \det[x_i^{2m-j}]_{1 \leq i, j \leq 2m}}{\det[x_i^{2m-j}]_{1 \leq i, j \leq 2m} \prod_{i=1}^m (x_{2i} - x_{2i-1})} = \frac{\det[x_i^{2m-j+\lambda_j}]_{1 \leq i, j \leq 2m}}{\prod_{j=1}^m (x_{2i} - x_{2i-1})}$$

as $x_{2i-1} \rightarrow x_{2i}$. Row operations on the matrix $[x_i^{2m-j+\lambda_j}]_{i, j=1, \dots, 2m}$ don't change the determinant, and subtraction of row $2i-1$ from $2i$ for $i = 1, 2, \dots, m$ yields

$$\frac{\det \begin{bmatrix} x_{2i}^{2m-j+\lambda_j} & x_{2i}^{2m-j+\lambda_j} - x_{2i-1}^{2m-j+\lambda_j} \\ & x_{2i-1}^{2m-j+\lambda_j} \end{bmatrix}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq 2m}}}{\prod_{j=1}^m (x_{2i} - x_{2i-1})} = \det \begin{bmatrix} x_{2i}^{2m-j+\lambda_j} & x_{2i-1}^{2m-j+\lambda_j} / (x_{2i} - x_{2i-1}) \\ & x_{2i-1}^{2m-1-j+\lambda_j} \end{bmatrix}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq 2m}}.$$

Using L'Hôpital's rule as $x_{2i-1} \rightarrow x_{2i}$ for $i = 1, \dots, m$, and then relabeling $x_{2i} \rightarrow x_i$, we get that the integrand in the theorem statement equals

$$\det \left[\begin{matrix} (2m-j+\lambda_j)x_i^{2m-1-j+\lambda_j} \\ x_i^{2m-j+\lambda_j} \end{matrix} \right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq 2m}}.$$

This determinant is a sum over the symmetric group S_{2m}

$$\sum_{\pi \in S_{2m}} \varepsilon(\pi) \prod_{i=1}^m (2m - \pi(2i) + \lambda_{\pi(2i)}) x_i^{2m-1-\pi(2i)+\lambda_{\pi(2j)}} x_i^{2m-\pi(2i-1)+\lambda_{\pi(2i-1)}},$$

where $\varepsilon(\pi)$ is the sign of π . If we require that $\pi(2j) > \pi(2j-1)$ for all j then the integrand becomes

$$\sum_{\pi \in S_{2m}, \pi(2j) > \pi(2j-1)} \varepsilon(\pi) \prod_{i=1}^m \left(((2m - \pi(2i) + \lambda_{\pi(2i)}) - (2m - \pi(2i-1) + \lambda_{\pi(2i-1)})) \right. \\ \left. x_i^{2m-1-\pi(2i)+\lambda_{\pi(2i)}} x_i^{2m-\pi(2i-1)+\lambda_{\pi(2i-1)}} \right)$$

and with the notation $y_i = 2m + \lambda_i - i$ it equals

$$\sum_{\pi \in S_{2m}, \pi(2j) > \pi(2j-1)} \varepsilon(\pi) \prod_{i=1}^m (y_{\pi(2i)} - y_{\pi(2i-1)}) x^{y_{\pi(2i)} + y_{\pi(2i-1)} - 1}$$

Integrating over each x_j separately gives the following value of the integral in the theorem:

$$\sum_{\pi \in S_{2m}, \pi(2j) > \pi(2j-1)} \varepsilon(\pi) \frac{y_{\pi(2i)} - y_{\pi(2i-1)}}{y_{\pi(2i)} + y_{\pi(2i-1)}}.$$

The sum is actually the Pfaffian of an anti symmetric matrix, in general for $A = [a_{i,j}]_{i,j=1,\dots,2m}$ anti symmetric

$$\text{Pf}[A] = \frac{1}{m!} \sum_{\pi \in S_{2m}, \pi(2j) > \pi(2j-1)} \varepsilon(\pi) \prod_{i=1}^m a_{\pi(2i-1), \pi(2i)}.$$

Hence

$$\int_{[0,1]^m} S_\lambda(x_1, x_1, x_2, x_2, \dots, x_m, x_m) \prod_{0 \leq i < j \leq m} (x_i - x_j)^4 \, d\mathbf{x} = m! \text{Pf} \left[\frac{y_i - y_j}{y_i + y_j} \right]_{1 \leq i, j \leq 2m},$$

but we know that in general this Pfaffian evaluates to the stated product formula, a result usually attributed to Schur [24].

The first odd case is true. For higher odd, start off as in the even but then expand the matrix along its column with only ones. This gives a sum of instances from the smaller even dimensional case, and by a degree argument the formula follows. \square

Corollary 3.6. *Let \mathbf{F} be a vector of k distinct graphs with edge counts $e_1 \geq \dots \geq e_k > 0$. Then the volume of the convex hull of the k -dimensional spine of \mathbf{F} is*

$$\text{Vol}(\text{conv} \{(p^{e_1}, p^{e_2}, \dots, p^{e_k}) \mid 0 \leq p \leq 1\}) = \frac{1}{k!} \prod_{1 \leq i < j \leq k} \frac{e_i - e_j}{e_i + e_j}.$$

Special cases of this integral have been computed before. Karlin and Shapley [14] used the Selberg integral formula [10, 25] to do the case of $e_i = i$, but they arrived to the integral in a completely different way. Selberg's formula can only be applied in the cases where the edges are of consecutive magnitudes: $e_i = e + i$.

3.2. Duality. As for polytopes there is a duality theory for convex hulls of algebraic sets [3]. The dual of the convex hull of the moment curve $\{(p, p^2, \dots, p^n) \mid 0 \leq p \leq 1\}$ parametrizes the degree n polynomials that are non-negative on the interval $[0, 1]$. The convex hulls of generalized moment curves are inside polytopes from subgraphs statistics, so the polytopes can be used to certify that polynomials are non-negative.

Proposition 3.7. *Let P be a polytope containing the generalized moment curve $\{(p^{e_1}, p^{e_2}, \dots, p^{e_d}) \mid 0 \leq p \leq 1\}$. If $\langle (c_1, c_2, \dots, c_d), v \rangle \geq -1$ for all vertices v of P then the polynomial $q(x) = 1 + c_1 x^{e_1} + c_2 x^{e_2} + \dots + c_d x^{e_d}$ is non-negative on the interval $[0, 1]$.*

Proof. The value of the polynomial $q(x) = 1 + c_1 x^{e_1} + c_2 x^{e_2} + \dots + c_d x^{e_d}$ at the point $x = p$ is $1 + \langle (p^{e_1}, p^{e_2}, \dots, p^{e_d}), (c_1, c_2, \dots, c_d) \rangle$. This implies that if $\langle (p^{e_1}, p^{e_2}, \dots, p^{e_d}), (c_1, c_2, \dots, c_d) \rangle \geq -1$ for all p in $[0, 1]$ then the polynomial q is positive on $[0, 1]$. \square

Example 3.8. Our running example $P_{(K_3, C_4, K_4 \setminus e); 6}$ in Figure 1 is perhaps not the most interesting polytope to certify non-negativity with, but we use it in an example anyways. The polynomial $p(x) = 1 - \frac{16}{3}x^3 + \frac{11}{2}x^4 - \frac{1}{2}x^5$ is non-negative on $[0, 1]$ since $(-\frac{16}{3}, \frac{11}{2}, -\frac{1}{2}) \cdot v \geq -1$ for all vertices v of $P_{(K_3, C_4, K_4 \setminus e); 6}$. Note that the point $(-\frac{16}{3}, \frac{11}{2}, -\frac{1}{2})$ is dual to the facet with vertices $(8/20, 10/45, 16/90)$, $(10/20, 15/45, 30/90)$, $(5/20, 3/45, 6/90)$, which one can find using Figure 2.

4. CURVY ZONOTOPES

In this section we generalize the ideas used to construct spines. A very general class of random graph models called exchangeable random graph models is used instead of the Erdős-Rényi model, but the idea is the same: The expectation values from random graph models give easily parameterized sets inside the polytopes of subgraph statistics, and these sets are useful to derive information about the polytopes. The exchangeable random graph models are related to the graph limits developed by Lovasz and Szegedy [15], as explained by Diaconis and Janson [6]. The machinery of graph limits is used more heavily in the next section where more of it is explained. In this section only the random graph models are needed.

The exchangeable random graph models is a generalization of the Erdős-Rényi random graph model $\mathcal{G}(n, p)$, obtained by replacing p by a measurable and symmetric function $W : [0, 1]^2 \rightarrow [0, 1]$. Denote the set of these measurable symmetric function by \mathcal{W} .

Definition 4.1. Let W be a function in \mathcal{W} . A graph G from the *exchangeable random graph model* $\mathcal{G}(n, W)$ is given as follows. The vertex set of G is $[n]$. Let X_1, \dots, X_n be independent random variables with uniform distribution on $[0, 1]$. For every pair $1 \leq i < j \leq n$ add an edge ij to G with probability $W(X_i, X_j)$.

Note that for given X_1, \dots, X_n the probability to get a graph with edge set E from $\mathcal{G}(n, W)$ is

$$\prod_{ij \in E} W(X_i, X_j) \prod_{ij \in E(K_n) \setminus E} (1 - W(X_i, X_j)).$$

We collect some basic facts from [15]. To proceed further the expectation values of subgraph densities is computed for the exchangeable graph models.

Proposition 4.2. *Let F be a graph with vertex set $[m]$ and let $W \in \mathcal{W}$. The expected value of $t(F, G)$ when G is from $\mathcal{G}(n, W)$ is*

$$\mathbb{E}(t(F, G)) = \int_{[0,1]^m} \prod_{ij \in E(F)} W(x_i, x_j) dx_1 \cdots dx_m,$$

where it is assumed that $m \leq n$

We use the notation $t(F, W) = \int_{[0,1]^m} \prod_{ij \in E(F)} W(x_i, x_j) dx_1 \cdots dx_m$ for the expectation values calculated in Proposition 4.2. For a vector of graphs $\mathbf{F} = (F_1, \dots, F_d)$ we define $t(\mathbf{F}, W) = (t(F_1, W), \dots, t(F_d, W))$.

In general it might be hard to compute $t(\mathbf{F}, W)$ explicitly.

Definition 4.3. A $W \in \mathcal{W}$ is a *stepfunction*, if for some partition of $[0, 1]$ into intervals, the value of $W(x, y)$ is completely determined by which parts of $[0, 1]$ that x and y belongs to.

For any symmetric $n \times n$ matrix M with entries in $[0, 1]$, the stepfunction W_M takes the value of M at row i and column j at $[(i-1)/n, i/n) \times [(j-1)/n, j/n)$.

Note that the value of W_M is not defined on the entire boundary of $[0, 1]^2$, this do not matter for our purposes since the boundary have measure zero.

The stepfunctions are useful since any function in \mathcal{W} can be approximated by a stepfunction, and for stepfunctions W it is possible to establish a polynomial expression for $t(\mathbf{F}, W)$.

Proposition 4.4. *Let $M = [m_{i,j}]_{n \times n}$ be a symmetric matrix with entries in $[0, 1]$ and let \mathbf{F} be a vector of graphs. Then*

$$t(\mathbf{F}, W_M) = (p_{F_1;n}(M), p_{F_2;n}(M), \dots, p_{F_d;n}(M)),$$

where

$$p_{F;n}(M) = \frac{1}{n^{|F|}} \sum_{\phi: V(F) \rightarrow [n]} \prod_{ij \in E(F)} m_{\phi(i), \phi(j)}.$$

Proof. Let F be a graph with vertex set $[k]$. Recall that the expectation value $t(F, W_M)$ is

$$\int_{[0,1]^k} \prod_{ij \in E(F)} W_M(x_i, x_j) dx_1 \cdots dx_k.$$

Partition $[0, 1]^k$ into the parts $[(i_1-1)/n, i_1/n) \times [(i_2-1)/n, i_2/n) \times \cdots \times [(i_k-1)/n, i_k/n)$, the boundary have measure zero and can be ignored. Observe that there is a bijection between the functions $\phi : V(F) \rightarrow [n]$ and the parts in the partition of $[0, 1]^k$. The

function ϕ corresponds to the part $[(\phi(1) - 1)/n, \phi(1)/n) \times [(\phi(2) - 1)/n, \phi(2)/n) \times \cdots \times [(\phi(k) - 1)/n, \phi(k)/n)$.

The function $\prod_{ij \in E(F)} W_M(x_i, x_j)$ is constant $\prod_{ij \in E(F)} m_{\phi(i), \phi(j)}$ on the part corresponding to ϕ . The integral of $\prod_{ij \in E(F)} W_M(x_i, x_j)$ over the part corresponding to ϕ is then $\frac{1}{n^{|F|}} \prod_{ij \in E(F)} m_{\phi(i), \phi(j)}$.

The integral over $[0, 1]^k$ then splits into the sum

$$\int_{[0, 1]^k} \prod_{ij \in E(F)} W_M(x_i, x_j) dx_1 \cdots dx_k = \frac{1}{n^{|F|}} \sum_{\phi: V(F) \rightarrow [n]} \prod_{ij \in E(F)} m_{\phi(i), \phi(j)}.$$

□

The polynomials $p_{F;n}(M)$ in Proposition 4.4 can be computed and evaluated, this makes it possible to study their image as a subset of the polytopes.

Definition 4.5. Let $\mathbf{F} = (F_1, \dots, F_d)$ be a vector of graphs and n a positive integer. The *curvy zonotope* is

$$Z_{\mathbf{F};n} = \left\{ (p_{F_1;n}(\mathbf{x}), p_{F_2;n}(\mathbf{x}), \dots, p_{F_d;n}(\mathbf{x})) \mid \mathbf{x} \in [0, 1]^{n^2}, x_{ij} = x_{ji} \right\}$$

where

$$p_{F;n}(\mathbf{x}) = \frac{1}{n^{|F|}} \sum_{\phi: V(F) \rightarrow [n]} \prod_{ij \in E(F)} x_{\phi(i)\phi(j)}.$$

The curvy zonotope $Z_{(K_3, C_4, K_4 \setminus e);2}$ is drawn in Figure 7.

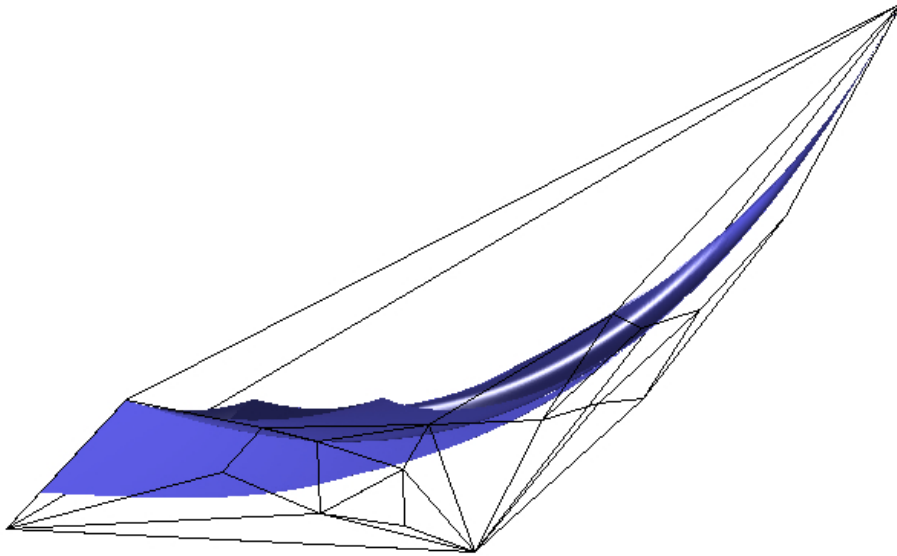


FIGURE 7. The curvy zonotope $Z_{(K_3, C_4, K_4 \setminus e);2}$. As defined in Definition 4.5.

The curvy zonotope $Z_{\mathbf{F};n}$ consists of the expected values from random graphs models according to Proposition 4.4, and this implies that it is a subset of the polytopes $P_{\mathbf{F};m}$. The curvy zonotopes contain the spines as the spines come from the special case of the

function W being constant p . The curvy zonotopes have many desirable properties that the spine do not, for example full dimensionality.

Theorem 4.6. *Let $\mathbf{F} = (F_1, \dots, F_d)$ be a vector of d distinct graphs with no isolated vertices and with $n_0 \geq \max\{|F_1|, \dots, |F_d|\}$. If $n \geq n_0$ then the curvy zonotope $Z_{\mathbf{F};n}$ is not contained in a hyperplane in \mathbb{R}^d .*

Proof. The curvy zonotope is defined in terms of polynomials coming from graphs. The idea of the proof is to identify a term in the polynomials with a special exponent. This special exponent ensure that no linear combination of the polynomials can be constant and this proves that the curvy zonotope is not contained in a hyperplane.

Let F be a graph in \mathbf{F} . The graph F give a polynomial

$$p_{F;n}(\mathbf{x}) \frac{1}{n^{|F|}} \sum_{\phi: V(F) \rightarrow [n]} \prod_{ij \in E(F)} x_{\phi(i)\phi(j)}$$

as in the definition of the curvy zonotope $Z_{\mathbf{F};n}$.

Assume that the curvy zonotope is in the hyperplane defined by $C_{|F_1|}p_{F_1;n}(\mathbf{x}) + \dots + C_{|F_d|}p_{F_d;n}(\mathbf{x}) = C$. Let F be a graph with maximal order among the graphs F_i with $C_{F_i} \neq 0$.

Let $\phi : V(F) \rightarrow [n]$ be injective. The function ϕ give a term $\frac{1}{n^{|F|}} \prod_{ij \in E(F)} x_{\phi(i)\phi(j)}$ in the polynomial $p_{F;n}(\mathbf{x})$. That ϕ is injective implies that the graph F' , with the indices occurring in $\prod_{ij \in E(F)} x_{\phi(i)\phi(j)}$ as vertex set and the pairs of indices occurring in $\prod_{ij \in E(F)} x_{\phi(i)\phi(j)}$ as edge set, is isomorphic to F .

Let F' be a graph with $C_{F'} \neq 0$ and let $p_{F';n}$ contain $\frac{1}{n^{|F|}} \prod_{ij \in E(F)} x_{\phi(i)\phi(j)}$. The function ϕ was injective from $V(F)$ and so $|F'|$ must be at least as large as $|F|$, but $|F|$ was maximal and hence $|F| = |F'|$. The function $\phi' : V(F') \rightarrow [n]$ giving the term $\frac{1}{n^{|F|}} \prod_{ij \in E(F)} x_{\phi(i)\phi(j)}$ is then injective. But it is possible to reconstruct F from $\frac{1}{n^{|F|}} \prod_{ij \in E(F)} x_{\phi(i)\phi(j)}$, assuming that the function giving the term is injective. This implies that no other polynomial from a graph F' with $C_{F'} \neq 0$ contains the term $\prod_{ij \in E(F)} x_{\phi(i)\phi(j)}$. This ensure that no term can cancel the unique term $C_F \frac{1}{n^{|F|}} \prod_{ij \in E(F)} x_{\phi(i)\phi(j)}$ in the polynomial $C_{|F_1|}p_{F_1;n}(\mathbf{x}) + \dots + C_{|F_d|}p_{F_d;n}(\mathbf{x})$, and so the polynomial can not be constant. This is a contradiction and the curvy zonotope is not contained in a hyperplane. \square

The curvy zonotope is not in a hyperplane, this means that there are points in it that span d -dimensional simplex and then all the polytopes $P_{\mathbf{F};n}$ contain this simplex and is full dimensional. This establishes the following corollary:

Corollary 4.7. *Let $\mathbf{F} = (F_1, \dots, F_d)$ be a vector of d distinct graphs with no isolated vertices and with $n_0 \geq \max\{|F_1|, \dots, |F_d|\}$. There is a constant $V_{\mathbf{F}} > 0$ such that the volume of $P_{\mathbf{F};n}$ is larger than $V_{\mathbf{F}}$, for $n \geq n_0$.*

The volume of the polytopes do decrease with increasing n . But it is not the case that the volume can become arbitrarily small.

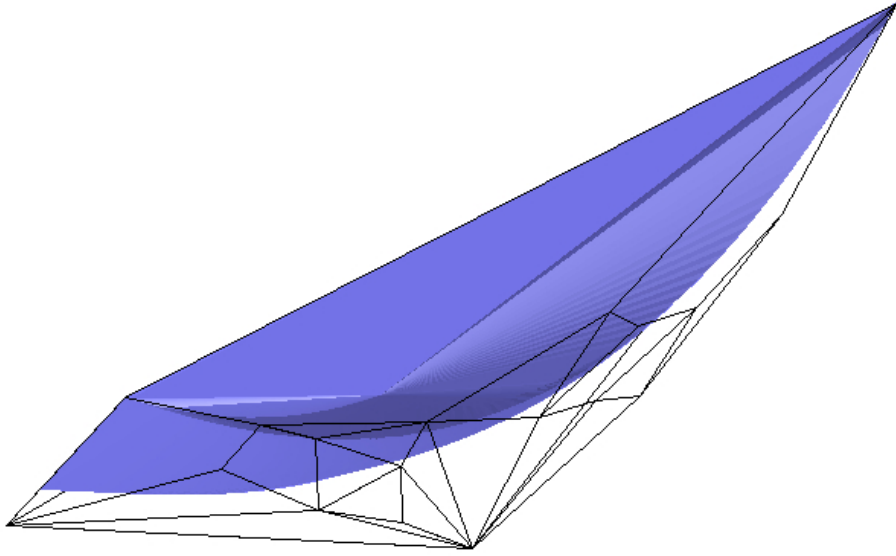


FIGURE 8.

5. LIMIT OBJECT

Recall that the polytopes of subgraph statistics satisfy the inclusion $P_{\mathbf{F},n} \subseteq P_{\mathbf{F},n_0}$ when $n \leq n_0$. This makes it possible to define a limit object.

Definition 5.1. Let \mathbf{F} be a vector of graphs where no entry have more than n_0 vertices. Define the *limit object*

$$P_{\mathbf{F};\infty} = \bigcap_{n \geq n_0} P_{\mathbf{F};n}.$$

Describing the polytopes is very hard for most n , but for very large n the polytopes should essentially look like the limit object. Finding good descriptions of the limit objects and their properties is the goal of this section. As it turns out, the objects are closely related to graph limits and the exchangeable random graph models.

It seems almost hopeless to get a simple description of the polytopes $P_{\mathbf{F},n}$ for any nontrivial vector \mathbf{F} and integer n . For the limit objects the situation is much better. Known results in extremal graph theory can be used to get good descriptions of the limit objects.

Razborov found in [19] the exact minimal possible triangle density for given edge density, when the number of vertices in the graph go to infinity. Razborov used machinery called flag algebras that he developed in [20] which is related to graph limits. If $T(p)$ is the minimal possible triangle density for edge density p , then $\{(p, T(p)) \mid p \in [0, 1]\}$ is a curve in the limit object $P_{(K_2, K_3), \infty}$. The convex hull of the curve $\{(p, T(p)) \mid p \in [0, 1]\}$ is the limit object $P_{(K_2, K_3), \infty}$. The convex hull of the curve has a nice explicit description: it's the hull of $\{(p, T(p)) \mid p \in [0, 1]\}$ is $\text{conv} \left\{ (0, 0), (1, 1), \left(1 - \frac{1}{2}, 0\right), \left(1 - \frac{1}{3}, \frac{2}{9}\right), \dots, \left(1 - \frac{1}{k}, \frac{(k-1)(k-2)}{k^2}\right), \dots \right\}$.

In this section the theory of graph limits is used to reinterpret the limit object, and to give a way to study it when an explicit description is not known.

In [15] limits of certain sequences of graphs were defined. A sequence of graphs $(G_i)_{i=1,2,\dots}$ is said to converge if the sequence $(t(F, G_i))_{i=1,2,\dots}$ converges for every graph F . One of the important facts about convergent sequences of graphs is that they look like they come from exchangeable random graph models. One result in this direction from [15] is the following:

Theorem 5.2 (Lovasz-Szegedy [15]). *If the sequence of graphs $(G_i)_{i=1,2,\dots}$ converges, then there is a $W \in \mathcal{W}$ such that $t(F, W) = \lim_{i \rightarrow \infty} t(F, G_i)$ for all graphs F . Let $W \in \mathcal{W}$ and let $(G_i)_{i=1,2,\dots}$ be a sequence of random graphs generated by the exchangeable random graph model $\mathcal{G}(i, W)$, then $(G_i)_{i=1,2,\dots}$ converges almost surely. In fact $t(F, W) = \lim_{i \rightarrow \infty} t(F, G_i)$ almost surely for all graphs F .*

The main ingredient in the proof is a method to pick a subsequence of $(G_i)_{i=1,2,\dots}$ with Szemerédi regularity properties. Using this subsequence it is possible to construct a suitable W .

There is a useful norm on the space of functions $[0, 1]^2 \rightarrow \mathbb{R}$.

Definition 5.3. The *rectangle norm* of a function $W : [0, 1]^2 \rightarrow \mathbb{R}$ is defined by

$$\|W\|_{\square} = \sup_{A \subseteq [0,1]^2} \int_A W(x_1, x_2) dx_1 dx_2.$$

In the previous section we calculated with stepfunctions. In more general situations it might be hard to get exact results, but it is possible to approximate any measurable by a stepfunction.

Lemma 5.4. *For any vector \mathbf{F} of d graphs with at most e edges, and any $W \in \mathcal{W}$, there is a stepfunction U with at most $2^{k_d \varepsilon^{-2}}$ steps (all of the same width) so that*

$$|t(\mathbf{F}, W) - t(\mathbf{F}, U)| < de\varepsilon.$$

Here k_d is a fixed constant depending on d .

Proof. This is essentially a straightforward application of the weak Szemerédi regularity lemma contained in [8], as explained in [15]. There is a constant k , so that for any $\varepsilon > 0$ and $W \in \mathcal{W}$, there exists a stepfunction U (with all steps of the same width) and at most $2^{k\varepsilon^{-2}}$ steps, so that $\|U - W\|_{\square} < \varepsilon$. The rectangle norm used, relates to the fact, that, according to Lemma 4.1 of [15], $|t(F, U) - t(F, W)| \leq |E(F)| \|U - W\|_{\square}$. \square

One part of the proof of Theorem 5.2 uses the following very useful result.

Lemma 5.5 (Lovasz-Szegedy [15]). *Let $(G_i)_{i \geq 1}$ be a sequence of graphs. Then there is an infinite subsequence $(G_i)_{i \geq 1, i \in S}$ and a $W \in \mathcal{W}$ so that*

$$\lim_{i \rightarrow \infty, i \in S} t(F, G_i)$$

exists for all graphs F , and it equals $t(F, W)$.

The limit object, the intersection of polytopes of subgraph statistics, is related to the graph limits of Lovasz and Szegedy. The limit object is the convex hull of expectation values.

Theorem 5.6. *For every vector \mathbf{F} of graphs, $P_{\mathbf{F}, \infty} = \text{conv} \{t(\mathbf{F}, W) \mid W \in \mathcal{W}\}$.*

Proof. Denote with $P_{\mathbf{F};\mathcal{W}}$ the convex hull defined in the theorem statement.

First we prove that any $x \in P_{\mathbf{F};\infty}$ is in $P_{\mathbf{F};\mathcal{W}}$. By definition $P_{\mathbf{F};\infty}$ is the intersection of the d -dimensional polytopes $P_{\mathbf{F};n}$ for n is larger than some n_0 . By triangulating the interior of every $P_{\mathbf{F};n}$ using only boundary vertices, we can describe x in a simplex for every $n > n_0$:

$$x = \sum_{j=0}^d \alpha_{j,n} v_{j,n}$$

where each $v_{j,n}$ is a vertex of $P_{\mathbf{F};n}$, each $\alpha_{j,n} \geq 0$, and $\sum_{j=0}^d \alpha_{j,n} = 1$. For every vertex in a $P_{\mathbf{F};n}$ there is a graph realizing that subgraph statistics. Choose $d+1$ graph sequences $(G_n^j)_{n>n_0}^{0 \leq j \leq d}$ so that $t(\mathbf{F}, G_n^j) = v_{j,n}$.

Use Lemma 5.5 on $(G_n^0)_{n>n_0}$ to find an index set S_0 and a symmetric measurable function $W_0 : [0, 1]^2 \rightarrow [0, 1]$ so that $\lim_{n \rightarrow \infty, n \in S_0} t(F, G_n^0)$ exists for all graphs F , and it equals $t(F, W_0)$.

For $j = 1, 2, \dots, d$ repeat the use Lemma 5.5, but on $(G_n^j)_{n \in S_{j-1}}$ to find an index set S_j and a symmetric measurable function $W_j : [0, 1]^2 \rightarrow [0, 1]$ so that $\lim_{n \rightarrow \infty, n \in S_j} t(F, G_n^j)$ exists for all graphs F , and it equals $t(F, W_j)$.

The points $t(\mathbf{F}, W_0), \dots, t(\mathbf{F}, W_d)$ spans (a possibly degenerate) simplex in $P_{\mathbf{F};\mathcal{W}}$ and we should locate x in it. The point $(\alpha_{0,n}, \alpha_{1,n}, \dots, \alpha_{d,n})$ is in a d -simplex for all n , and hence there is a subsequence $S' \subset S_d$ so that $\lim_{n \rightarrow \infty, n \in S'} \alpha_{j,n} = \tilde{\alpha}_j$.

Collecting the preceding limits, gives the desired

$$x = \sum_{j=0}^d \tilde{\alpha}_j t(\mathbf{F}, W_j).$$

Now we prove the other direction. Pick a point x in $P_{\mathbf{F};\mathcal{W}}$ parametrized by

$$x = \sum_{j=0}^d \alpha_j t(\mathbf{F}, W_j)$$

and assume that $x \notin P_{\mathbf{F};\infty}$. Since $P_{\mathbf{F};\infty}$ is a closed convex set there is a hyperplane H strictly separating x and $P_{\mathbf{F};\infty}$. By definition $P_{\mathbf{F};\infty}$ is the intersection of $P_{\mathbf{F};n_0} \supset P_{\mathbf{F};n_0+1} \supset P_{\mathbf{F};n_0+2} \supset \dots$ for some n_0 .

Assume that all polytopes $P_{\mathbf{F};n}$ intersect the hyperplane H , and pick a convergent sequence $\{x_i\}_{i=n_0, n_0+1, \dots}$ such that $x_i \in P_{\mathbf{F};i} \cap H$. The sequence $\{x_i\}_{i=n_0, n_0+1, \dots}$ converge to something in H and the polytopes are compact, this implies that the limit of $\{x_i\}_{i=n_0, n_0+1, \dots}$ is in all polytopes $P_{\mathbf{F};i}$. This is a contradiction, H will not intersect all $P_{\mathbf{F};n}$. For n_1 large enough, the hyperplane H will not intersect $P_{\mathbf{F};n_1}$, and x and $P_{\mathbf{F};n_1}$ are separated by H .

Generate random graphs $G_n^j \in \mathcal{G}(n, W_j)$ for $n \geq n_1$. According to Lemma 5.5, $\lim_{n \rightarrow \infty} t(\mathbf{F}, G_n^j)$ exists and converges to $t(\mathbf{F}, W_j)$. Note that all of the points $t(\mathbf{F}, G_n^j)$ are in the closed convex set $P_{\mathbf{F};n_1}$, and hence so is all their limits $t(\mathbf{F}, W_j)$. But then $x = \sum_{j=0}^d \alpha_j t(\mathbf{F}, W_j)$ should also be in $P_{\mathbf{F};n_1}$, which contradicts that x and $P_{\mathbf{F};n_1}$ are separated by a hyperplane. \square

The curvy zonotopes can also be used to approximate the limit object arbitrary well.

Theorem 5.7. *Let \mathbf{F} be a vector of d graphs on at most e edges. If $\varepsilon > 0$ and $n > 2^{k_d \frac{d^2 e^2}{\varepsilon^2}}$ then*

$$0 \leq \text{Vol}(P_{\mathbf{F};\infty}) - \text{Vol}(\text{conv } Z_{\mathbf{F};n}) < \varepsilon.$$

Here k_d is a constant depending only on d .

Proof. Any point in the limit object $P_{\mathbf{F};\infty}$ is in the convex hull of

$$\{t(\mathbf{F}, W) \mid W : [0, 1]^2 \rightarrow [0, 1] \text{ symmetric and measurable}\},$$

by Theorem 5.6, and the points of $Z_{\mathbf{F};n}$ is a subset of those, since they are all on the form $t(\mathbf{F}, U)$ for some stepfunction U with n steps. According to Lemma 5.4 any $t(\mathbf{F}, W)$ is within distance $k_d^{-1}\varepsilon$ of a point in the curvy zonotope $Z_{\mathbf{F};n}$, and then according to Lemma 3.2 we are done. \square

6. CONJECTURES ABOUT THE LIMIT OBJECT $P_{\mathbf{F};\infty}$.

We have previously inscribed cyclic polytopes in the limit objects. We will now define another cyclic polytope and conjecture that a particular class of limit objects actually are cyclic polytopes.

The vertices of $P_{(K_2, K_n);\infty}$ are given by the limits of complete k -equipartite graphs according to results by Bollobas [4, 5]. It is not hard to see that $P_{(K_2, K_n);\infty}$ is a cyclic polytope, and we believe that this is true in a more general setting.

For positive integers $e_1 < e_2 < \dots < e_m$ define the *tail* $s^e : [0, 1] \rightarrow [0, 1]^m$ by

$$s_i^e(x) = \prod_{j=1}^{e_i-1} (1 - jx).$$

Proposition 6.1. *The convex hull of any finite set of points on a tail is a cyclic polytope.*

Proof. Let $e_i = i + 1$. Consider the matrix whose columns are the points on the polytope, ordered by increasing x . The polynomial on row i is of degree i and has nonzero coefficients on all terms of lower degree. There is then a sequence of row operations that takes this matrix to a matrix with the same determinant but the polynomial on row i is a monomial of degree i . In this situation the argument from Ziegler's book [28] applies and the polytope is cyclic.

The general case, with arbitrary e_i is obtained from the case $e_i = i + 1$ by projection: any tail is obtained from the tail with $e_i = i + 1$ by removing some coordinates. The tail with $e_i = i + 1$ is generic and this is the property which ensure that points on it span cyclic polytopes, removing coordinates from a generic curve give a new generic curve. \square

Conjecture 6.2. *Let $e_1 < e_2 < \dots < e_m$ be positive integers and s^e their tail. The convex hull of $\mathbf{1}$ and $\{s^e(1/k) \mid k = 1, 2, 3, \dots\}$ is $P_{(K_{e_1}, K_{e_2}, \dots, K_{e_m});\infty}$.*

The conjectured vertex description of $P_{(K_{e_1}, K_{e_2}, \dots, K_{e_m});\infty}$ also gives a facet description since it's essentially a cyclic polytope. If we would chop off the vertex $\mathbf{1}$ from the convex hull described in Conjecture 6.2 with an hyperplane, then the remaining convex set would be an ordinary polytope. We believe this is true in the following general form.

Conjecture 6.3. *For any vector \mathbf{F} of graphs there is a positive integer m , such that for any $\varepsilon > 0$, the limit object $P_{\mathbf{F};\infty}$ can be chopped down to a polytope with a finite number of vertices, by using m hyperplanes to remove at most a volume ε .*

ACKNOWLEDGEMENT

In a manuscript of this article we calculated several special cases of an integral and conjectured that there was an easy combinatorial description. This was proved by Peter Forrester and we are very happy that he allows us to include his proof in Theorem 3.5.

Alexander Engström is a Miller Research Fellow, and gratefully acknowledges support from the Miller Institute for Basic Research in Science at UC Berkeley.

Patrik Norén gratefully acknowledges support from the Wallenberg foundation.

REFERENCES

- [1] David J. Aldous. Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.* **11** (1981), no. 4, 581–598.
- [2] David J. Aldous. Exchangeability and Continuum Limits of Discrete Random Structures, in *Proceedings of the International Congress of Mathematicians Hyderabad, India, 2010* Vol. I, World Scientific Publishing Company, 2010.
- [3] Alexander Barvinok. *A course in convexity*. Graduate Studies in Mathematics, 54. American Mathematical Society, Providence, RI, 2002. 366 pp.
- [4] Béla Bollobás. Relations between sets of complete subgraphs. *Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen, 1975)*, pp. 79–84. *Congr. Numer.* No. XV, Utilitas Math., Winnipeg, Man., 1976.
- [5] Béla Bollobás. *Extremal graph theory*. Reprint of the 1978 original. *Dover Publications, Inc., Mineola, NY*, 2004. 488 pp.
- [6] Persi Diaconis, Svante Janson. Graph limits and exchangeable random graphs. *Rend. Mat. Appl.* (7) **28** (2008), no. 1, 33–61.
- [7] Reinhard Diestel. *Graph theory*. Fourth edition. Graduate Texts in Mathematics, 173. Springer, Heidelberg, 2010. 437 pp.
- [8] Alan Frieze, Ravi Kannan. Quick approximation to matrices and applications. *Combinatorica* **19** (1999), no. 2, 175–220.
- [9] Peter J. Forrester. *Log-gases and random matrices*. London Mathematical Society Monographs Series, 34. Princeton University Press, Princeton, NJ, 2010. 791 pp.
- [10] Peter J. Forrester, Ole S. Warnaar. The importance of the Selberg integral. *Bull. Amer. Math. Soc. (N.S.)* **45** (2008), no. 4, 489–534.
- [11] Lars Hörmander. *The analysis of linear partial differential operators. I*. Distribution theory and Fourier analysis. Reprint of the second (1990) edition. Classics in Mathematics. Springer-Verlag, Berlin, 2003. 440 pp.
- [12] Douglas N. Hoover. Relations on probability spaces and arrays of random variables. Institute of Advanced Study, Princeton, preprint 1979.
- [13] Olav Kallenberg, *Probabilistic symmetries and invariance principles*. Probability and its Applications. Springer, New York, 2005. 510 pp.
- [14] Samuel Karlin, Lloyd S. Shapley. Geometry of moment spaces. *Mem. Amer. Math. Soc.*, 1953, no. 12, 93 pp.
- [15] László Lovász, Balázs Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** (2006), no. 6, 933–957.
- [16] Madan Lal Mehta. *Random matrices*. Third edition. Pure and Applied Mathematics (Amsterdam), 142. Elsevier/Academic Press, Amsterdam, 2004. 688 pp.
- [17] Ezra Miller, Bernd Sturmfels. *Combinatorial commutative algebra*. Graduate Texts in Mathematics, 227. Springer-Verlag, New York, 2005. 417 pp.

- [18] Vladimir Nikiforov. The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.* **363** (2011), 1599–1618.
- [19] Alexander A. Razborov. On the minimal density of triangles in graphs. *Combin. Probab. Comput.* **17** (2008), no. 4, 603–618.
- [20] Alexander A. Razborov. Flag algebras. *J. Symbolic Logic* **72** (2007), no. 4, 1239–1282.
- [21] Alessandro Rinaldo, Stephen E. Fienberg, Yi Zhou. On the geometry of discrete exponential families with application to exponential random graph models. *Electron. J. Stat.* **3** (2009), 446–484.
- [22] Garry Robins, Pip Pattison, Yuval Kalish, Dean Lusher. An Introduction to Exponential Random Graph (p^*) Models for Social Networks. *Social Networks* **29** (2007), no. 2, 173–191.
- [23] Bruce E. Sagan. *The symmetric group. Representations, combinatorial algorithms, and symmetric functions*. Second edition. Graduate Texts in Mathematics, 203. Springer-Verlag, New York, 2001. 238 pp.
- [24] Issai Schur. Über die Darstellung der symmetrischen und der alternierenden Gruppe durch gebrochene lineare Substitutionen, *J. Reine Angew. Math.* **139** (1911), 155–250.
- [25] Atle Selberg. Bemerkinger om et multipelt integral, *Norsk Mat. Tidsskr.* **26** (1944), 71–78.
- [26] Kristian Ranestad, Bernd Sturmfels. On the convex hull of a space curve. [arXiv:0912.2986](#), preprint 2009, 18 pp.
- [27] Kristian Ranestad, Bernd Sturmfels. The convex hull of a variety. [arXiv:1004.3018](#), preprint 2010, 12 pp.
- [28] Günter M. Ziegler. *Lectures on polytopes*. Graduate Texts in Mathematics, 152. Springer-Verlag, New York, 1995. 370 pp.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720

E-mail address: alex@math.berkeley.edu

KTH – THE ROYAL INSTITUTE OF TECHNOLOGY, STOCKHOLM, SWEDEN

E-mail address: pnore@math.kth.se